

**An Expectation-Maximization  
Approach to Spectral Structure  
Identification and Estimation**

Marian Viola, Alan Bolton and  
Bill Moran

DSTO-RR-0207

**DISTRIBUTION STATEMENT A**  
Approved for Public Release  
Distribution Unlimited

20020509 045

# **An Expectation-Maximization Approach to Spectral Structure Identification and Estimation**

***Marian Viola, Alan Bolton and Bill Moran***

**Surveillance Systems Division  
Electronics and Surveillance Research Laboratory**

**DSTO-RR-0207**

## **ABSTRACT**

The method of Expectation-Maximization is applied to find an approximate maximum-likelihood estimate of the amplitudes, phases and frequency shift of a known collection of spectral lines in a signal with Gaussian noise. The technique works well for many frequency components provided that these frequencies are well separated. It is applied to the identification of the frequencies present in a radar return from a propeller driven aircraft, arising from modulation by blades of the propellers.

**APPROVED FOR PUBLIC RELEASE**



*AQ F02-08-1537*

*Published by*

*DSTO Electronics and Surveillance Research Laboratory  
PO Box 1500  
Salisbury, South Australia, Australia 5108*

*Telephone: (08) 8259 5555  
Facsimile: (08) 8259 6567*

*© Commonwealth of Australia 2001  
AR No. AR 011-791  
October, 2001*

**APPROVED FOR PUBLIC RELEASE**

## **An Expectation-Maximization Approach to Spectral Structure Identification and Estimation**

### **EXECUTIVE SUMMARY**

Many problems in signal processing require the estimation of the significant frequency components of a signal; that is, of the frequency, amplitude and phase. In particular such problems arise in estimating the spectral content of returns from aircraft where discrete frequency components may arise from modulation by the engines. Good estimates of these components are useful in identification of the aircraft type from the radar return.

Here is presented a method of estimation of discrete frequency (line) components where the inter-spacing is known but the absolute positions of the lines are unknown, as would be the case when an aircraft return had unknown doppler. Also unknown are the amplitude and phase of the lines. These parameters are estimated by a technique based on the Expectation-Maximization algorithm which, iteratively, produces good approximations to the maximum likelihood estimate.

The method is found to work extremely well with synthetic data with up to six line components and large amounts of added noise. When tested on real signals, it is able to handle up to 21 line components.



## Authors

### **Marian Viola**

*Marian Viola is a Principal Research Scientist and the Head of the Radar Target Identification Discipline of Surveillance Systems Division. He has a B.E. from the University of Adelaide. He has worked on the design and development of radar components and measurement radar systems, specializing in coherent signal generation. His research interests include radar non-cooperative target recognition, radar waveform design and LPI radar.*

---

### **Alan Bolton**

*Alan Bolton is a Senior Research Scientist in the Surveillance Systems Division of DSTO specializing in target identification. He has worked as a lecturer in Electrical Engineering both at the South Australian Institute of Technology and at the University of Queensland. He has a B.E. and a PhD in Electrical Engineering from the University of Adelaide. His areas of expertise are control systems and signal processing.*

---

### **Bill Moran**

*Bill Moran has been a Professor of Electrical Engineering at the University of Melbourne since July 2001. Previously was Professor of Mathematics at Flinders University (1991–2001) and at the University of Adelaide (1976–1991).*

*He has a B.Sc. (Hons) degree in Mathematics from the University of Birmingham in the UK and a Ph.D. from the University of Sheffield. He was elected a Fellow of the Australian Academy of Science in 1984. From 1991 to 1998, he was a Program Leader in the Cooperative Research Centre for Sensor Signal and Information Processing.*

*His research areas in mathematics include harmonic analysis, representation theory and number theory. More recently he has added research interest in several aspects of signal processing. He is CEO of Eudoxus Consulting, through which he does much of his current work with DSTO.*

---

DSTO-RR-0207

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Description of Basic Algorithm</b>	<b>2</b>
<b>3</b>	<b>Choice of <math>\sigma_r^2</math></b>	<b>5</b>
<b>4</b>	<b>Simulation and Results</b>	<b>5</b>
<b>5</b>	<b>Test on Real Data</b>	<b>9</b>
<b>6</b>	<b>Conclusion</b>	<b>12</b>
	<b>References</b>	<b>12</b>

## Figures

<b>1</b>	<b>Magnitude of simulated data . . . . .</b>	<b>6</b>
<b>2</b>	<b>Convergence of parameter estimates using normalizing method . . . . .</b>	<b>7</b>
<b>3</b>	<b>Convergence of parameter estimates using annealing method . . . . .</b>	<b>8</b>
<b>4</b>	<b>Real Data: Frequency Domain . . . . .</b>	<b>9</b>
<b>5</b>	<b>Convergence of parameter estimates using normalizing algorithm . . . . .</b>	<b>10</b>
<b>6</b>	<b>Convergence of parameter estimates using annealing algorithm . . . . .</b>	<b>11</b>

DSTO-RR-0207

# 1 Introduction

In many radar signal processing problems a signal with known spectral structure needs to be identified and its other parameters such as amplitudes and phases of the individual lines and an overall frequency shift (due to doppler) estimated. The aim of this paper is to describe an algorithm for extracting this information from a signal given a knowledge of the relative frequencies of its line components.

Radar returns from propeller driven aircraft contain signal components due to reflections from the blades of the propellers. These components are manifested, in part, by modulation of the reflected *rf* signal by a number of discrete frequencies. At long ranges, the radar returns themselves are significantly contaminated by noise. Consequently, standard signal processing methods have not been easily able to extract the frequency components and give good estimates of their phases and amplitudes.

The underlying frequency components in returns from a propeller driven aircraft are modified both by the effects of doppler and by variations in the shaft rate of the engine. The former, under the narrow band approximation, which is valid here, shifts the frequencies while the latter scales them. Both of these effects need to be taken into account in any method that attempts to identify aircraft by these means. However for the purposes of this paper we restrict attention only to frequency shift, leaving scaling to a future publication.

We describe here a method based on the Expectation-Maximization (EM) algorithm of Dempster *et. al.*, ([1]), for extracting from a signal the amplitudes and phases of a given set of frequency components as well as their frequency offset. The elaboration of the algorithm in [2] is closely followed.

The method then, begins with the description of a signal model comprising a small number of pure frequency components, but of unknown phase and amplitude. In addition the frequencies are only known relative to each other and not absolutely. White noise of unknown variance is added to this signal. To be specific, we assume a complex signal  $s(t)$  with known frequencies  $f_1, f_2, \dots, f_R$ , an unknown frequency shift  $\tau$ , unknown phases  $\phi_1, \phi_2, \dots, \phi_R$ , and unknown amplitudes  $\theta_1, \theta_2, \dots, \theta_R$ . Thus the signal is

$$s(t) = \sum_{r=1}^R \theta_r e^{2\pi i ((f_r + \tau)t + \phi_r)} + w(t) \quad (1)$$

where  $w(t)$  is complex Gaussian white noise with variance  $\sigma^2$ , which is also assumed unknown. We write

$$\begin{aligned} \underline{\theta} &= (\theta_1, \theta_2, \dots, \theta_R), \\ \underline{\phi} &= (\phi_1, \phi_2, \dots, \phi_R), \\ \mathbf{f} &= (f_1, f_2, \dots, f_R). \end{aligned}$$

Our aim, as already indicated, is to estimate the various unknown parameters of the signal from a data record,  $S = (s(k))_{k=1}^M$ . A standard maximum-likelihood approach appears not to be practical, as there is no closed form solution to the optimization problem. Instead a modified form of the EM algorithm is used. The EM algorithm is known, in many situations, to provide good convergence to the maximum likelihood estimate.

In this report we show how the EM algorithm is applied to the problem described above and obtain results for simulated data with up to 6 different frequency components. Two variations on the EM technique with slightly different update rules in the iteration process are examined. Both of these are found to work well with relatively high levels of noise. In addition we show how the algorithm works with a typical return from a propeller driven aircraft. There reasonable estimates for over 20 frequency components are obtained.

## 2 Description of Basic Algorithm

The probability density of a record  $S = (s(k))_{k=1}^M$  generated by a signal model of the form (1) is

$$p_{\underline{\theta}, \underline{\phi}, \tau}(S) = \frac{1}{(\sqrt{2\pi\sigma^2})^M} \exp\left(-\frac{1}{2\sigma^2} \sum_{k=1}^M \left| s(k) - \sum_{r=1}^R \theta_r e^{2\pi i ((f_r + \tau)k + \phi_r)} \right|^2\right). \quad (2)$$

From this we observe that the salient terms in the log-likelihood of this record are

$$L_s(\underline{\theta}, \underline{\phi}, \tau) = -\frac{M}{2}\sigma^2 - \frac{1}{2\sigma^2} \sum_{k=1}^M \left| s(k) - \sum_{r=1}^R \theta_r e^{2\pi i ((f_r + \tau)k + \phi_r)} \right|^2. \quad (3)$$

Note first that, once  $\underline{\theta}$ ,  $\underline{\phi}$ ,  $\tau$  are known, then an estimate for  $\sigma^2$  can easily be found by maximizing the log-likelihood (3):

$$\widehat{\sigma^2} = \left( \frac{1}{M} \sum_{k=1}^M \left| s(k) - \sum_{r=1}^R \theta_r e^{2\pi i ((f_r + \tau)k + \phi_r)} \right|^2 \right)^{1/2}. \quad (4)$$

In fact, estimates for  $\sigma^2$  do not enter into the algorithm except at this point. Accordingly this aspect of the problem will be ignored from now on.

To find the estimates of the other parameters the EM approach as described in [2] is followed. To simplify formulae, note that the non-constant part of the log-likelihood is (the negative of)

$$Q_s(\underline{\theta}, \underline{\phi}, \tau) = \sum_{k=1}^M \left| s(k) - \sum_{r=1}^R \theta_r e^{2\pi i ((f_r + \tau)k + \phi_r)} \right|^2.$$

The key ingredient in the EM approach is the introduction of new (*unobservable*) random variables

$$\mathbf{y}_k^{(r)} = \theta_r e^{2\pi i ((f_r + \tau)k + \phi_r)} + \mathbf{w}_k^{(r)}$$

where the random variables  $\mathbf{w}_k^{(r)}$  are complex Gaussian of mean zero and variance  $\sigma_r^2$ . The variances  $\sigma_r^2$  are chosen to sum to  $\sigma^2$ , so that

$$\mathbf{s}(k) = \sum_{r=1}^M \mathbf{y}_k^{(r)}.$$

We shall say more later about how these  $\sigma_r^2$  are chosen. The appropriate terms in the log-likelihood for this vector of random variables are then

$$G_{\mathbf{Y}}(\underline{\theta}, \underline{\phi}, \tau) = \sum_{r=1}^R \frac{1}{\sigma_r^2} \sum_{k=1}^M |y_k^r - \theta_r e^{2\pi i ((f_r + \tau)k + \phi_r)}|^2.$$

The “expectation” phase of the EM algorithm is accommodated by calculating the expectation of this random variable conditioned on the random vector  $\mathbf{S} = (\mathbf{s}(k))_{k=1}^M$ , where parameter values  $\underline{\theta}^\sharp, \underline{\phi}^\sharp, \tau^\sharp$  are the values assumed in the distribution for  $\mathbf{S}$ . The particular values of  $\underline{\theta}, \underline{\phi}, \tau$  that minimize this expression are found and are used as updated versions of  $\underline{\theta}^\sharp, \underline{\phi}^\sharp, \tau^\sharp$  in an iterative scheme.

The conditional expectation operator with respect to  $\mathbf{S}$  under the distribution with the parameters  $\underline{\theta}^\sharp, \underline{\phi}^\sharp, \tau^\sharp$  is denoted by  $\mathcal{E}^\sharp$  and the absolute expectation by  $\mathcal{E}$ . Thus we need to consider

$$\mathcal{E}^\sharp(G_{\mathbf{Y}}(\underline{\theta}, \underline{\phi}, \tau)) = \sum_{r=1}^R \frac{1}{\sigma_r^2} \sum_{k=1}^M \left( \mathcal{E}^\sharp(|y_k^r|^2) - 2\Re(\theta_r e^{2\pi i ((f_r + \tau)k + \phi_r)} \mathcal{E}^\sharp(y_k^r)^*) + \theta_r^2 \right),$$

where  $*$  denotes complex conjugation. The first term under the inner summation sign does not vary with the parameters so we may ignore it in the maximization process. Our aim, then, is to find  $\underline{\theta}, \underline{\phi}, \tau$  to maximize

$$U(\underline{\theta}, \underline{\phi}, \tau) = \sum_{r=1}^R \left( \frac{1}{\sigma_r^2} \sum_{k=1}^M \Re(\theta_r e^{2\pi i ((f_r + \tau)k + \phi_r)} \mathcal{E}^\sharp(y_k^r)^*) - M\theta_r^2 \right). \quad (5)$$

To calculate  $\mathcal{E}^\sharp(y_k^r)$ , note that the random variables involved are Gaussian and so

$$\mathcal{E}^\sharp(y_k^r) = \mathcal{E}(y_k^r) + \mathbf{C}_{ys} \mathbf{C}_{ss}^{-1} (\mathbf{S} - \mathcal{E}(\mathbf{S}))_{k,r} \quad (6)$$

where  $\mathbf{C}_{ys}$  is the covariance of  $\mathbf{Y}$  with  $\mathbf{S}$  and  $\mathbf{C}_{ss}$  is the auto-covariance of  $\mathbf{S}$ . The latter is just  $\sigma^2 \mathbf{I}$ , and the  $((k, r), k')$  component of the former is 0 unless  $k = k'$ ; when  $k = k'$  it equals  $\sigma_r^2$ . This gives

$$\mathcal{E}^\sharp(y_k^r)^* = \theta_r^\sharp e^{-2\pi i ((f_r + \tau^\sharp)k + \phi_r^\sharp)} + \frac{\sigma_r^2}{\sigma^2} \left( \mathbf{s}(k)^* - \sum_{q=1}^R \theta_q^\sharp e^{-2\pi i ((f_q + \tau^\sharp)k + \phi_q^\sharp)} \right).$$

Substituting in (6), we find that

$$U(\underline{\theta}, \underline{\phi}, \tau) = \sum_{r=1}^R \left( \frac{1}{\sigma_r^2} \sum_{k=1}^M \Re(\theta_r e^{2\pi i ((f_r + \tau)k + \phi_r)} (\theta_r^\sharp e^{-2\pi i ((f_r + \tau^\sharp)k + \phi_r^\sharp)}) + \right. \quad (7)$$

$$\left. \frac{\sigma_r^2}{\sigma^2} \left( \mathbf{s}(k)^* - \sum_{q=1}^R \theta_q^\sharp e^{-2\pi i ((f_q + \tau^\sharp)k + \phi_q^\sharp)} \right) \right) - M\theta_r^2 \Bigg). \quad (8)$$

In order to deal with this expression, we define

$$a_{k,r} = \theta_r^\sharp e^{-2\pi i ((f_r + \tau^\sharp)k + \phi_r^\sharp)} + \frac{\sigma_r^2}{\sigma^2} \left( \mathbf{s}(k)^* - \sum_{q=1}^R \theta_q^\sharp e^{-2\pi i ((f_q + \tau^\sharp)k + \phi_q^\sharp)} \right), \quad (9)$$

and let  $A$  be the matrix  $(a_{k,r})$ . Notice that all of this expression is calculable in terms of an instantiation  $S = (s(k))$  of the random vector  $\mathbf{S}$  and the previous values of the parameters.

Let

$$c_{k,r} = e^{2\pi i((f_r + \tau)k + \phi_r)},$$

and let  $C$  denote the matrix  $(c_{k,r})$ . Then (7) can be rewritten as

$$U(\underline{\theta}, \underline{\phi}, \tau) = \sum_{r=1}^R \frac{\theta_r}{\sigma_r^2} \Re(C^* A)_{r,r} - M \frac{\theta_r^2}{2\sigma_r^2}, \quad (10)$$

where now  $*$  denotes the Hermitian transpose.

Once  $\underline{\phi}$  and  $\tau$  are known then it is easy to maximize  $U$  for  $\theta_r$  by choosing

$$\theta_r = \frac{1}{M} (C^* A)_{r,r}. \quad (11)$$

The values of  $\underline{\phi}$  and  $\tau$  are obtained by maximizing

$$\sum_{r=1}^R \frac{\theta_r}{\sigma_r^2} \Re(C^* A)_{r,r}.$$

In fact we use the prior estimate  $\theta_r^\sharp$  rather than  $\theta_r$  here. In a sense this departs from the formalism of the EM-algorithm, however the effect does not seem to be significant and simulations indicate that the method converges in practice. It is difficult, in this case, to see how the strict methodology of the EM-algorithm could be followed. Thus we maximize

$$\sum_{r=1}^R \frac{\theta_r^\sharp}{\sigma_r^2} \Re(C^* A)_{r,r}. \quad (12)$$

Let us write

$$d_{k,r} = a_{k,r} e^{-2\pi i f_r k}$$

so that (12) becomes

$$\sum_{r=1}^R \frac{\theta_r}{\sigma_r^2} \sum_{k=1}^M d_{k,r} e^{-2\pi i (\tau k + \phi_r)}. \quad (13)$$

Given the correct choice of  $\tau$ , the correct choice of  $\phi_r$  to maximize this are

$$\phi_r = \arg \left( \sum_{k=1}^M d_{k,r} e^{-2\pi i \tau k} \right). \quad (14)$$

For this choice of  $\phi_r$  the sum (13) becomes

$$\sum_{r=1}^R \frac{\theta_r}{\sigma_r^2} \left| \sum_{k=1}^M d_{k,r} e^{-2\pi i \tau k} \right|^2$$

The problem then is to maximize this over  $\tau$ . Once this is done the previous equations (14), (11) can be used to update the remaining parameters.

The maximization over  $\tau$  is just an exhaustive search over a range of possible values of  $\tau$ . Since it is anticipated that the range of such values will be relatively small this can be accomplished quickly and easily. In the real example we use for illustration, an initial crude guess at the frequency shift was made by observing the FFT. This limited the range of  $\tau$  to be searched.

### 3 Choice of $\sigma_r^2$

In the algorithm presented in the previous section, the only constraint imposed in the choice of the variances  $\sigma_r^2$  of the virtual random variables  $y_k^r$  is that they sum to  $\sigma^2$ . To implement the algorithm a choice of these parameters is necessary.

We have experimented with three different choices. The first and most obvious choice is to make the  $\sigma_r^2$  all equal (and therefore equal to  $\sigma^2/R$ ). In fact this gave slow and often poor convergence results compared with the other two methods, especially when the values of the magnitudes  $\theta_r$  varied significantly. Such behaviour is not unexpected. If some of the  $\theta_r$  are small then this choice of  $\sigma_r^2$  imposes a comparatively large noise on the  $r$ th signal component. Because of its poor performance we do not report on simulations with this choice of  $\sigma_r^2$ .

Before describing the second choice of  $\sigma_r^2$ , we first consider a definition of  $\sigma_r^2$  (denoted by  $\sigma_{tr}^2$ ) that involves making them proportional to the magnitude of the corresponding signal component, that is,

$$\frac{\sigma_{tr}^2}{\sigma^2} = \frac{\theta_r'}{(\sum_{r=1}^R \theta_r'^2)^{1/2}}.$$

This, of course, satisfies the sum constraint. In fact it was found that using this method gave rise to problems with the components of very small magnitude. For these, the amount of adjustment imposed by the algorithm from one iteration to the next was too small and so a lower threshold was imposed on the size of the  $\sigma_r^2$ . To be precise, for the second choice of the variances, we define  $\sigma_{ur}^2$  by

$$\frac{\sigma_{ur}^2}{\sigma^2} = \max(\sigma_{tr}, 0.2 * (\sum_{r=1}^R \sigma_r^2)^{1/2}),$$

and then normalize to obtain  $\sigma_r^2$ :

$$\frac{\sigma_r^2}{\sigma^2} = \frac{\sigma_{ur}^2}{\sum_{r=1}^R \sigma_{ur}^2}.$$

As we shall see in the next section this method performs quite well both for simulated and real data. We call it the *normalizing* method.

A third choice, which works surprisingly well, requires relaxation of the constraint on the sum of the  $\sigma_r^2$ 's. In fact the values of  $\sigma_r^2$  are all equal at each iteration and are of the form

$$\sigma_r^2 = \frac{\sigma^2}{R} (1 + \log(1 + \rho)),$$

where  $\rho$  is the number of iterations remaining. This resembles the cooling schedule of optimization algorithms such as simulated annealing. We refer to this as the *annealing* method.

### 4 Simulation and Results

The algorithm works well on simulated data comprising up to 6 frequencies, as we shall illustrate with a particular example. In this example, the length of a data record is 1024, the frequencies used are

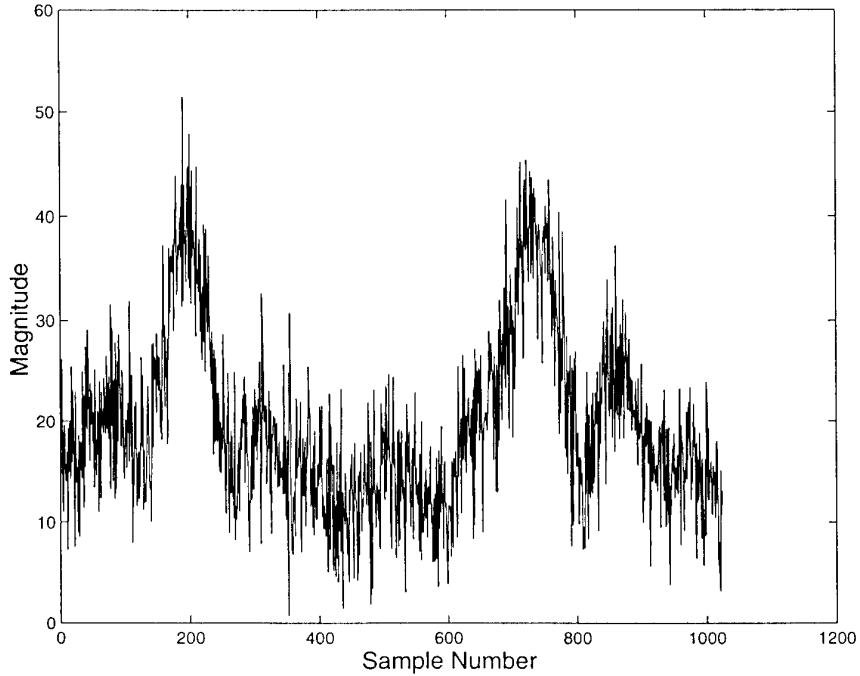
$$\mathbf{f} = (f_r)_{r=1}^6 = (1.1, 3.2, 4.8, 6.8, 8.9, 12.4),$$

and the corresponding magnitudes and phases

$$\underline{\theta} = (\theta_r)_{r=1}^6 = (6, 15, 8, 2, 7, 5),$$

$$\underline{\phi} = (\phi_r)_{r=1}^6 = (0.25, 0.8, 0.4, 0.3, 0.6, 0.7).$$

The frequency offset  $\tau$  was chosen to be equal to 0.7. Gaussian noise with standard deviation 7 was added to this signal. Figure 1 shows the magnitude of the signal with noise present. The

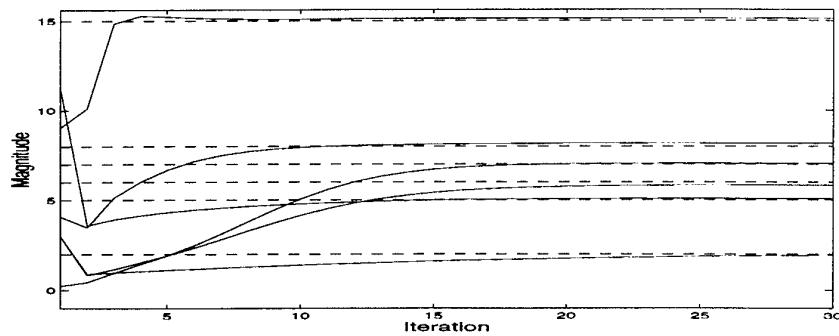


*Figure 1: Magnitude of simulated data*

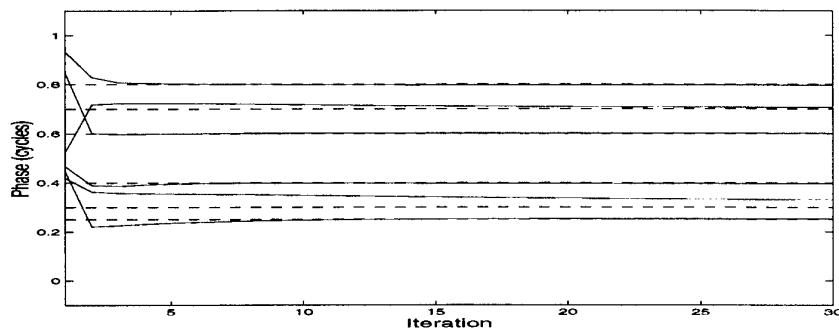
algorithm was started with a random choice of each of the parameters to be estimated and run for 50 iterations. The next two figures show the convergence of the various parameters towards the actual values (given by the dashed lines), first for the normalizing method (Figure 2) and then for the annealing method (Figure 3).

Both methods show good convergence. The normalizing method converged faster on the component with largest amplitude but less quickly on the remaining amplitudes. It also tended to give slightly less accurate answers for both the amplitudes and phases. In both cases, but especially in the normalizing case, the phase estimation was less accurate when the magnitude of that component was smaller. Using both methods the phase and frequency offset estimates tended to converge rapidly.

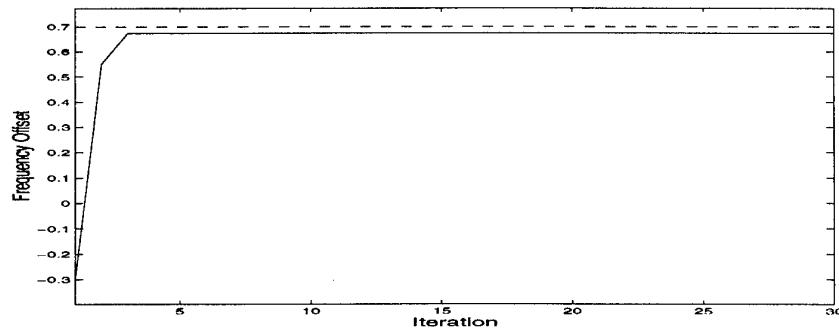
It was found that the algorithm worked better when the frequencies were well spaced. The presence of two close frequencies would, at least, slow the convergence and sometimes even result in instability of the algorithm. As a result the algorithm did not perform well on simulated data of this length with more than about 6 frequencies.



(a) Magnitude



(b) Phase



(c) Frequency Offset

Figure 2: Convergence of parameter estimates using normalizing method

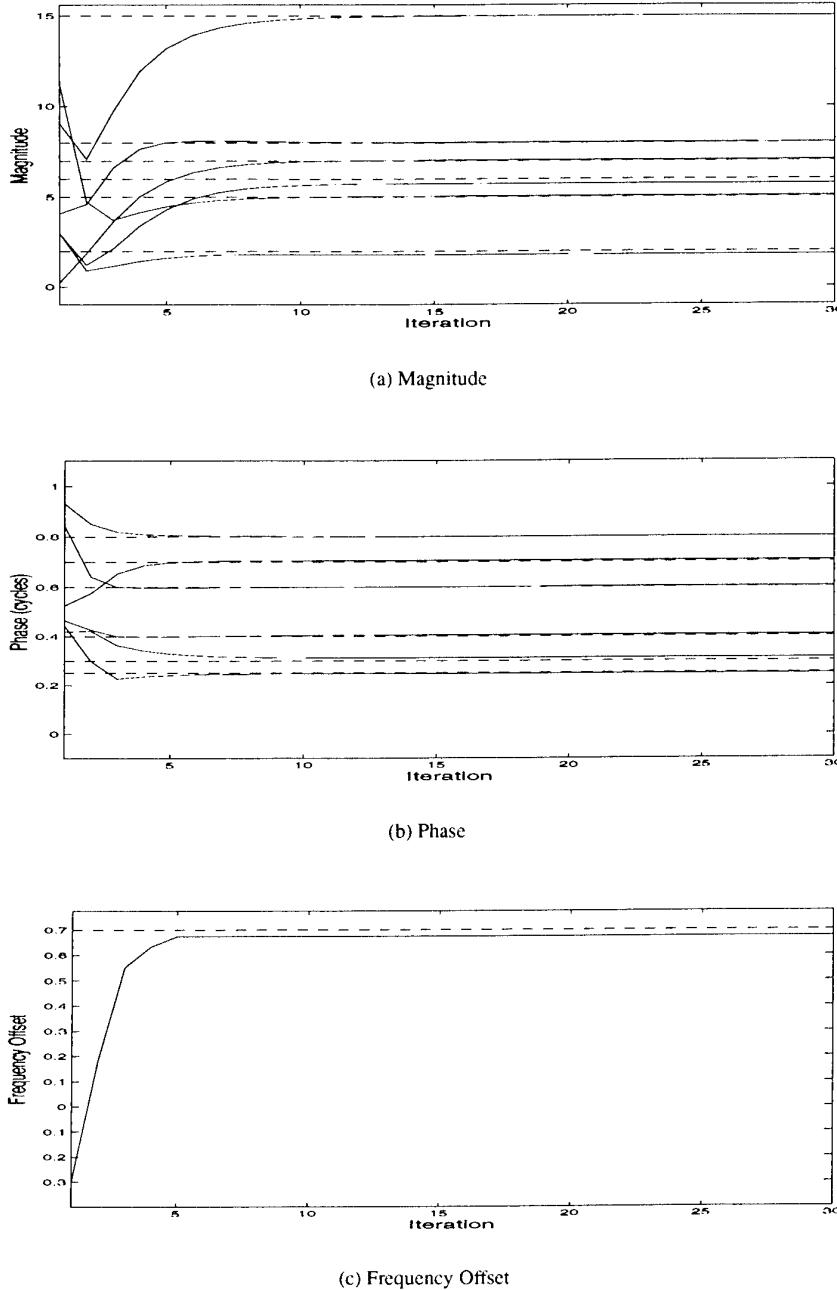
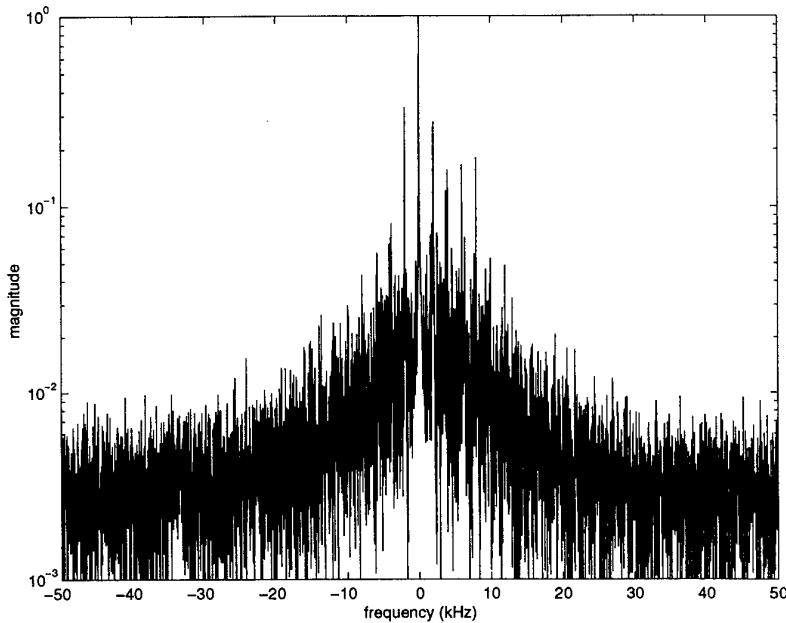


Figure 3: Convergence of parameter estimates using annealing method

The algorithm was tested with four frequencies and varying amounts of noise up to a standard deviation of about 100. While the estimates decreased in accuracy as the noise level increased, the algorithm continued to converge and to provide estimates that remained useful.

## 5 Test on Real Data

We now demonstrate the capability of the algorithm on some real radar returns from a propeller driven aircraft. Figure 4 gives the magnitude of the FFT of the original time domain data.



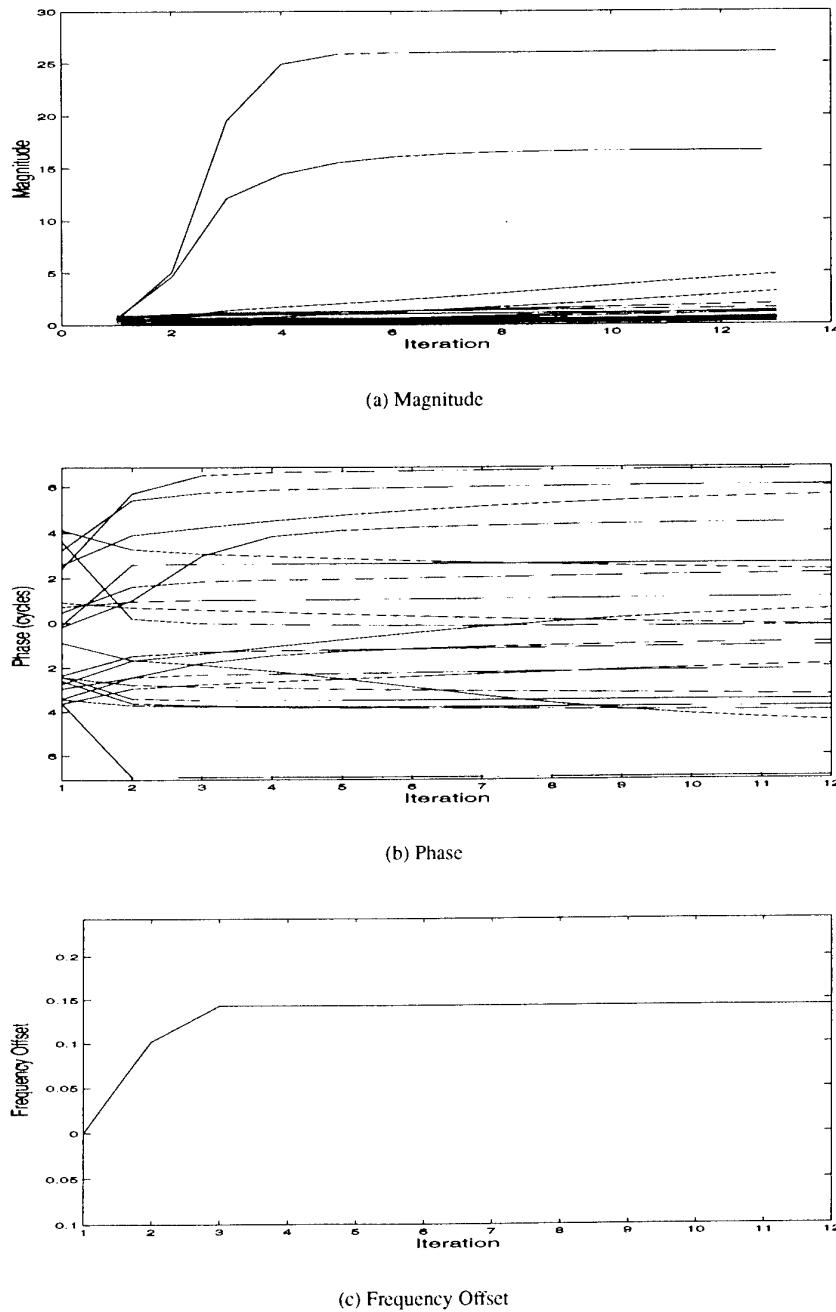
*Figure 4: Real Data: Frequency Domain*

As can be seen, the data contains a number of significant frequency components centred close to 427. We chose this number as an initial guess of the frequency shift so as to reduce the search range for  $\tau$ . Thus the initial choice of frequencies were 9 equally spaced components, centred on 427 with spacings of 28.31:

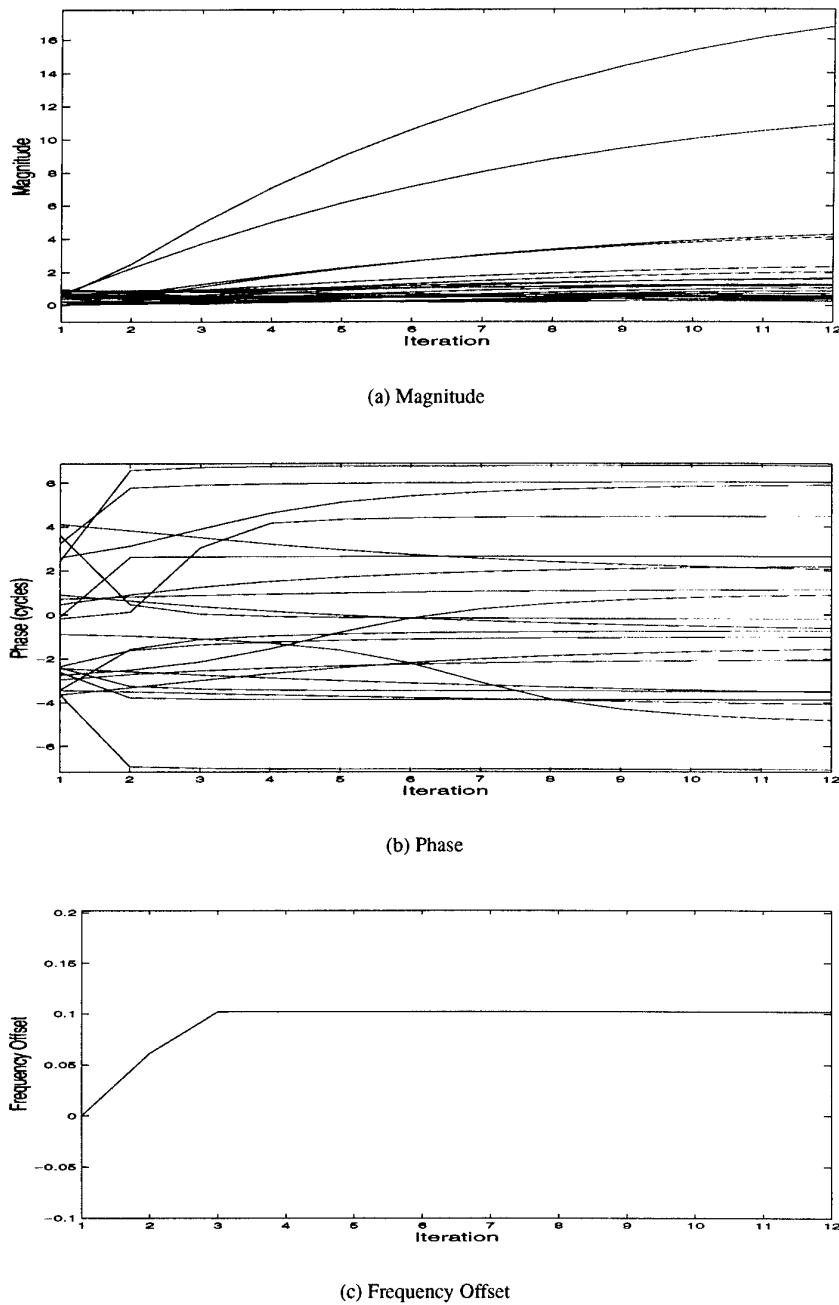
$$\mathbf{f} = (f_r) = (-4, -3, -2, -1, 0, 1, 2, 3, 4) \times 28.31 + 427.$$

It should be noted that, in this data, there was no prior knowledge of the doppler shift, so that, to test the accuracy of the method, further frequency shifts were imposed. The data used to provide the convergence graphs had a shift of 1.52 over the initial data.

Figure 5 shows the convergence of the various parameters over 12 iterations of the normalizing form of the algorithm. The annealing form of the algorithm was also applied to the data for 12 iterations. It yielded the results illustrated in Figure 6.



*Figure 5: Convergence of parameter estimates using normalizing algorithm*



*Figure 6: Convergence of parameter estimates using annealing algorithm*

Both algorithms yielded good results; the annealing method converging more slowly, in this case, than the normalizing one. What is, perhaps, significant here is that in this data there are many other significant frequency components, so that the data no longer conforms even crudely to the model. Nonetheless it was able using both methods to give a good estimate for the various parameters. As already stated we tested the algorithm with various artificial frequency offsets imposed. The results showed remarkable consistency, agreeing to within 0.006. We have obtained similar results with as many as 21 frequency components in this data.

## 6 Conclusion

We have described two versions of a method of estimating the magnitude, phase and frequency offset of a number of spectral lines of known relative frequency in a signal with Gaussian noise. The method is based on the expectation-maximization algorithm, and the two versions correspond to different ways of assigning the variances of the unobservable random variables.

Both methods work well for simulated data comprising up to six frequencies in relatively high levels of noise, and for over 20 frequencies from a complicated real signal with many line components in noise. It is difficult to compare the relative merits of the two methods of choosing the variances of the unobservable random variables. As a crude generalization, the annealing method slightly outperformed the normalizing method in speed of convergence while the normalizing method gave slightly more accurate estimates. However the results were very data dependent. The convergence rates for individual lines were highly dependent on the magnitude of the line for both methods. For the annealing method, convergence of the lines of larger magnitude was slower but for the smaller lines faster. The normalizing method was exactly the opposite.

Both methods worked less well with narrowly separated lines. We intend to report on an improved algorithm to handle this case at a later stage.

## References

1. A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Ann. Roy. Statist. Soc.*, vol. 11, pp. 1-38, 1977.
2. S. M. Kay, *Fundamentals of Statistical Signal Processing*. Prentice-Hall, 1993.

## DISTRIBUTION LIST

An Expectation-Maximization Approach to Spectral Structure Identification and Estimation  
 Marian Viola, Alan Bolton and Bill Moran

	Number of Copies
<b>DEFENCE ORGANISATION</b>	
<b>Task Sponsor</b>	
DGAD	1
<b>S&amp;T Program</b>	
Chief Defence Scientist	1
FAS Science Policy	1
AS Science Corporate Management	1
Director General Science Policy Development	1
Counsellor, Defence Science, London	Doc Data Sht
Counsellor, Defence Science, Washington	Doc Data Sht
Scientific Adviser to MRDC, Thailand	Doc Data Sht
Scientific Adviser Joint	1
Navy Scientific Adviser	Doc Data Sht
Scientific Adviser, Army	Doc Data Sht
Air Force Scientific Adviser	1
Director Trials	1
<b>Aeronautical and Maritime Research Laboratory</b>	
Director, Aeronautical and Maritime Research Laboratory	1
<b>Electronics and Surveillance Research Laboratory</b>	
Director, Electronics and Surveillance Research Laboratory	Doc Data Sht
Chief, Surveillance Systems Division	1
Research Leader, Tactical Surveillance Radar	1
Head, Radar Target Identification	1
Mr M Viola	1
Dr A Bolton	1
Prof W Moran, Contractor, SSD	1
Mr R Melino	1
<b>DSTO Research Library and Archives</b>	
Library Fishermans Bend	1
Library Maribyrnong	Doc Data Sht
Library Salisbury	1
Australian Archives	1

Library, MOD, Pyrmont		Doc Data Sht
Library, MOD, HMAS Stirling	1	
US Defense Technical Information Center	2	
UK Defence Research Information Centre	2	
Canada Defence Scientific Information Service	1	
NZ Defence Information Centre	1	
National Library of Australia	1	
<b>Capability Systems Staff</b>		
Director General Maritime Development		Doc Data Sht
Director General Land Development	1	
Director General Aerospace Development		Doc Data Sht
<b>Navy</b>		
SOSCI, Surface Combatants FEG Wartime Division, Maritime HQ, Garden Island		Doc Data Sht
<b>Army</b>		
ABCA Standardisation Officer, Tobruk Barracks, Puckapunyal 3662	4	
<b>Knowledge Staff</b>		
Director General Command, Control, Communications and Computers (DGC4)		Doc Data Sht
Director General Intelligence, Surveillance, Reconnaissance and Electronic Warfare		Doc Data Sht
Director General Defence Knowledge Improvement Team		Doc Data Sht
<b>Intelligence Program</b>		
DGSTA, Defence Intelligence Organisation	1	
Manager, Information Centre, Defence Intelligence Organisation	1	
<b>Acquisitions Program</b>		
<b>Corporate Support Program</b>		
Library Manager, DLS-Canberra	1	
<b>UNIVERSITIES AND COLLEGES</b>		
Australian Defence Force Academy Library	1	
Head of Aerospace and Mechanical Engineering, ADFA	1	
Deakin University Library, Serials Section (M List)	1	
Hargrave Library, Monash University		Doc Data Sht
Librarian, Flinders University	1	
<b>OTHER ORGANISATIONS</b>		
NASA (Canberra)	1	

---

AusInfo	1
State Library of South Australia	1

#### **ABSTRACTING AND INFORMATION ORGANISATIONS**

Engineering Societies Library, US	1
Documents Librarian, The Center for Research Libraries, US	1

#### **INFORMATION EXCHANGE AGREEMENT PARTNERS**

Acquisitions Unit, Science Reference and Information Service, UK	1
Library – Exchange Desk, National Institute of Standards and Technology, US	1
National Aerospace Laboratory, Japan	1
National Aerospace Laboratory, Netherlands	1

#### **SPARES**

DSTO Salisbury Research Library	5
---------------------------------	---

**Total number of copies:** 50

Page classification: UNCLASSIFIED

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA		1. CAVEAT/PRIVACY MARKING		
2. TITLE  An Expectation-Maximization Approach to Spectral Structure Identification and Estimation		3. SECURITY CLASSIFICATION  Document (U) Title (U) Abstract (U)		
4. AUTHORS  Marian Viola, Alan Bolton and Bill Moran		5. CORPORATE AUTHOR  Electronics and Surveillance Research Laboratory PO Box 1500 Salisbury, South Australia, Australia 5108		
6a. DSTO NUMBER DSTO-RR-0207	6b. AR NUMBER AR 011-791	6c. TYPE OF REPORT Research Report	7. DOCUMENT DATE October, 2001	
8. FILE NUMBER Z1111-222-33	9. TASK NUMBER AIR 87/113	10. SPONSOR DGAD	11. No OF PAGES 12	12. No OF REFS 2
13. URL OF ELECTRONIC VERSION  <a href="http://www.dsto.defence.gov.au/corporate/reports/DSTO-RR-0207.pdf">http://www.dsto.defence.gov.au/corporate/ reports/DSTO-RR-0207.pdf</a>		14. RELEASE AUTHORITY  Chief, Surveillance Systems Division		
15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT  <i>Approved For Public Release</i>  OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, SALISBURY, SOUTH AUSTRALIA 5108				
16. DELIBERATE ANNOUNCEMENT  No Limitations				
17. CITATION IN OTHER DOCUMENTS  No Limitations				
18. DEFTEST DESCRIPTORS  Radar signals Signal processing Spectral analysis				
19. ABSTRACT  The method of Expectation-Maximization is applied to find an approximate maximum-likelihood estimate of the amplitudes, phases and frequency shift of a known collection of spectral lines in a signal with Gaussian noise. The technique works well for many frequency components provided that these frequencies are well separated. It is applied to the identification of the frequencies present in a radar return from a propeller driven aircraft, arising from modulation by blades of the propellers.				

Page classification: UNCLASSIFIED

**RESEARCH REPORT DSTO-RR-0207 AR-011-791 OCTOBER 2001**

---



ELECTRONICS AND SURVEILLANCE RESEARCH LABORATORY  
PO BOX 1500 EDINBURGH SOUTH AUSTRALIA, 5111  
AUSTRALIA, TELEPHONE (08) 8259 5555